DOCUMENT RESUME

ED 225 383 FL 013 432

AUTHOR Perkins, Kyle
TITLE The Application of a Stylistic Metric to English as a Second Language Compositions.
PUB DATE [82]
NOTE 19p.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Adults; *English (Second Language); Evaluation Methods; *Grammar; Literary Styles; *Readability Formulas; Reading Comprehension; Semantics; *Writing (Composition)

ABSTRACT

A formula for assessing style is described and applied to compositions written by adult students of English as a second language. Style is defined as the total effect of what is done by a writer in the composing process, including choice of words, organization, and sentence and paragraph strategies. Earlier attempts to quantify style are reviewed. The formula proposed by Williams (1979) posits that the clearest style is that in which agents and subjects coincide. It is asserted that, in English, an agent and what it does semantically are manifested redundantly by at least three grammatical structures: the grammatically defined roles of subject and predicate, the sequence of subject-agent followed by what it does, and the form class of subject-agent as a noun and what it does as a verb. Williams' metric was applied to 10 compositions and the resulting ratings were compared with ratings obtained by other means. The results indicated that Williams' metric has concurrent validity for second language compositions. (RW)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED225383

THE APPLICATION OF A STYLISTIC METRIC TO ENGLISH AS A SECOND LANGUAGE COMPOSITIONS

Kyle Perkins

Summary

A recently proposed metric to assess style was applied to compositions written by adult students of English as a Second Language. The metric, proposed by Joseph Williams, rests on the proposition that "the clearest style is one in which the grammatical structures of a sentence most redundantly support the perceived semantic structure." There is evidence that the metric has concurrent validity in English as a Second Language composition.

FL 013 432

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Kyle Perkins

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.
Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

There is a robust tradition of attempting to identify and quantify the constructs, components and traits of simple and complex styles in native English speaker composition (referred to as L1 hereafter). This paper reports an attempt to apply a metric designed to assess style to compositions written by adult students of English as a Second Language (referred to as L2 hereafter). The metric employed in this research was proposed by Joseph Williams of the University of Chicago and reported in Williams (1979).

A preliminary point. For the purposes of this paper, style is assumed to be the total effect of all things a writer does in the composing process and, those things include, at minimum, the following: 1) choice of words, 2) organization, 3) sentence and paragraph strategies, 4) thought. This working definition is similar to that offered by Irmscher (1979) who defines style as "an aggregate of lexical and structural features of a particular work."

Attempts to identify, quantify and relate style to comprehension and comprehensibility are as varied as the number of active researchers in the field. For example, Rudolf Flesch is well-known for his readability formulae which establish a scale of simple to complex styles based on a relationship between the number of words per sentence and the number of syllables per word (cf Flesch 1948, 1962a and b). Flesch (1948) reported correlation coefficients of .66 and .47, respectively,

between the comprehension difficulty of texts and the average number of syllables and words in those texts.

For his work in programming computers to generate grammatical English sentences, Yngve (1960) proposed the "depth hypothesis." It is hypothesized that keeping track of various syntactic relationships loads a reader's short-term memory and taxes comprehension; the greater the number of syntactic relationships per word, the more difficult the sentence containing that word is to comprehend. The following cogent example of Yngve's hypothesis was offered by Williams; also see Figure 1:

PUT FIGURE 1 ABOUT HERE

> In the first, each successive word requires that we continue to store the fact that we must subsequently close the whole phrase. Each word postpones the moment when that will happen, and so we are required to hold in our short term memory an increasing number of grammatical bits. The other does not require us to anticipate and store nearly as many grammatical bits and to that degree, it is claimed, is easier for us to process (Williams 1979:599).

Mehler (1963) suggested a metric that depends on the number of transformational rules that apply to a deep structure to derive a surface structure. This approach no longer enjoys universal acceptance because of the enormous disagreement amongst theoretical linguists about the linguistic and cognitive nature of underlying representations and transformational rules.

It has been proposed that the number of T-units in a

writing sample be quantified and related to style. The T-unit is an index of syntactic complexity and was developed by Hunt (1965) who defines it as follows: "A main clause plus all its subordinate clauses and non-clausal structures attached to or embedded in it" (Hunt 1970:4). The idea behind the T-unit is that writers produce longer T-units, as they mature syntactically and that mature sentences are equated with stylistic complexity.

But, as Bartholomae (1980) -- discussing the problems of basic writers -- points out, "syntactic immaturity" is not always the cause of breakdown in written communication:

> When we have used conventional T-unit analysis, and included in our tabulations figures on words/clause, words/T-unit and clauses/T-unit that were drawn from 'intended T-units' as well as actual T-units, we have found that basic writers do not, in general, write 'immature' sentences. They are not, that is, 13th graders writing 7th grade sentences. In fact, they often attempt syntax whose surface is more complex than that of more successful freshman writers. They get into trouble by getting in over their heads, not only attempting to do more than they can, but imagining as their target a syntax that is more complex than convention requires (Bartholomae 1980:254).

In in an application of Yngve's depth hypothesis, Bormuth (1968) found that mean word depth, as well as frequency and length of independent clauses correlated significantly with comprehension difficulty. Bormuth determined the mean word depth for each independent clause in his test passages by averaging the depths assigned to the words within them and for

each passage by averaging the depths of the words in the passage.

Winterowd (1975) offered a format for stylistic analysis which includes calculating the average number of T-units/sentence, the average number of cumulative modifiers/T-unit and the percentages of kinds of cumulative modifiers used:

Sentences
- Arrangements
  - Loose and periodic
  - Passive
  - Other arrangements
- Types (simple, compound, complex, compound-complex)
- Cumulative modifiers
- Miscellaneous embeddings

Words (nouns, verbs, adjectives, adverbs)

Figures
- Metaphors
- Other figurative language

(Winterowd 1975:439)

For further examples of analyzing style along statistical dimensions, see Gibson (1966).

Hirsch (1977) -- proposing the concept of "relative readability" -- noted that "one prose style is better than another when it communicates the same meanings as the other does but requires less effort from the reader" [to reach semantic closure] .

A more recent metric to assess style is that proposed by Williams:

> The clearest style is one in which the grammatical structures of a sentence most redundantly support the perceived

> semantic structure. The more consistently the grammatical structure reinforces -- or reflects -- the semantic structure, the more easily a reader takes up that semantic structure ... -- the way semantic and grammatical structures do or do not coincide -- would appear to be a very salient metric of a complex style, if we agree to define complex in a special way -- as the style in which grammatical structures systematically fail to support semantic structures and thereby make sentences relatively more difficult for us to read. The features that create this kind of style are frequent in modern prose; they are crucial to the way we process sentences; they are structures that we teach our students to avoid; and we can demonstrate what happens when they do not. So as a metric of complex and simple styles, it meets every one of our criteria (Williams 1979: 603-604).

Williams finds support for his metric in the literature of linguistics, psycholinguistics, and speech perception and production. Williams cites Schlesinger's (1977) work that suggests that a reader assumes the subject of a sentence coincides with the agent or source of what the verb refers to. In addition, there is the suggestion that the English reader metaphorically transforms instruments into agents. Williams states, "we tentatively assume that all subjects are agents, unless some aspect of semantic or syntactic structure signals otherwise."

Williams' hypothesis is that a reader can more easily read prose (or a style) in which agents and subjects, verbs and what those agents do, coincide. His linguistic basis for this hypothesis is that subjects and agents tend to coincide

across different languages (cf Fillmore, 1968; Clark, 1965; Johnson, 1967; Fodor, Bever and Garrett, 1974). The manner in which subjects, verbs and semantic structures coincide is obviously language-specific. Williams claims that, in English, an agent and what it does semantically are manifested redundantly (in a communication science sense) by three grammatical structures, at minimum:

> (1) by the grammatically defined roles of subject and what is predicated of it;
>
> (2) by the sequence of subject-agent first and what it does second;
>
> (3) by form class: the subject/agent is a noun and what it does is a verb (Williams 1979:602).

In sum, Williams' metric seems to rest on the assumption that surface structures have a causative effect on how sentences are processed and sentences in which the grammatical structures redundantly support the perceived semantic structure are easier to process than those in which the redundancy between syntax and semantics is missing.

There is a potential flaw in Williams' metric in that he assumes that, generally speaking, there is a one-to-one relationship between the surface structure of a sentence and its meaning. This is not necessarily the case, because a particular meaning can be manifested by several sentences or paraphrases, and the opposite case is that in which a sentence might be

ambiguous or have several different meanings.

Consider the following sentences.

1. Max gave a Pelikan to Bruno.
2. Max gave Bruno a Pelikan.
3. A Pelikan was given to Bruno by Max.
4. Bruno was given a Pelikan by Max.
5. It was Max who gave Bruno a Pelikan.
6. Bruno was the one to whom Max gave a Pelikan.
7. It was a Pelikan that Max gave to Bruno.
8. It was Bruno to whom Max gave a Pelikan. etc.

It is assumed that 1-8 have the same meaning -- though the focus may differ in each -- and that 1-8 are derived from the same underlying representation shown in Figure 2. The different variations result from different transformations, e.g., dative movement, passive, etc., having been applied to the underlying representation.

PUT FIGURE 2 ABOUT HERE

Conversely, one surface structure may represent different meanings, which, by necessity, must be reflected by different deep structures or underlying representations. For example, fondling gorillas can be dangerous has at least two different meanings shown by the rudimentary underlying representations in Figure 3.

PUT FIGURE 3 ABOUT HERE

The previous discussion notwithstanding, Williams' metric

is one of the most promising to come along, because it is based on research findings from a broad spectrum of disciplines. No empirical data resulting from the application of Williams' metric could be found for L1 composition so the researcher decided to apply it to compositions written by L2 students.

The research question addressed in this paper is not whether a continuum of simple to complex styles could be isolated for a sample of L2 compositions, but whether a measure from Williams' metric would correlate with other sorts of evaluations, that is, would Williams' metric have any concurrent validity with other measures of attained L2 writing proficiency.

A small sample (n=10) of compositions written by students in a basic composition course for foreign students was randomly selected for analysis. Each composition was approximately 350 words long and written under test conditions. Each sentence was examined to determine whether its grammatical structure reflected the perceived semantic structure; a stylistic complexity quantity was calculated for each paper, figured as a percentage of sentences in which the grammatical structure correctly and adequately reflected or reinforced the semantic structure. Furthermore, the papers were given analytical ratings on fluency, grammar, mechanics, relevance and vocabulary by a team of independent raters using a 5-point scale (1-5), and the following descriptors were used for the rating scales: "fluency refers to readability, organization and co-

herence; grammar refers to cohesive markers, the complexity of sentence structures and sentence boundaries (e.g., no fragments or run-ons; mechanics include minor grammar rules (e.g., subject-verb agreement), spelling and punctuation; relevance means adherence to the topic, thesis development with supports and overall unity; vocabulary refers to the level of word choice, variety of vocabulary, appropriate usage and idioms."

The stylistic complexity data and the analytical rating data were submitted to the Spearman rank-order correlation procedure. When the significance of a rho is tested, the null hypothesis states that the correlation between the X and Y variables is zero in the population under study. The correlation may be either positive or negative. Because no empirical data of this sort had been analyzed before, the researcher was not able to predict a positive or negative correlation; therefore, a two-tailed test was seen as being appropriate. A two-tailed region of rejection at the .05 level of significance for this research is stated as follows:

R: rho $\geq$ .683, df 9

The rank-order correlations obtained are reported in Table 1.

PUT TABLE 1 ABOUT HERE

As Table 1 shows, the rank-order correlation between Williams' stylistic metric and the analytical rating on fluency almost reached significance at the .05 level; the obtained rho was .6750 but it would have to have been .683 or greater to

have been significant. The fact that each of the analytical variables correlated differently with Williams' metric is due to the tendency of raters to weight a few traits of writing differently.

The level of significance of the rho's is not as important for this study as the magnitude of the rho's. There appears to be overlap between Williams' metric of stylistic complexity and the analytical ratings of fluency and grammar. To the extent that the variables tend to covary, they tend to measure the same construct. That no single correlation went as high as .8 or .9 is not surprising for the following reasons: style is not a primary trait like concord which is readily amenable to quantification; style is an amalgam of many traits and factors and is the product of the interrelatedness of these traits/factors.

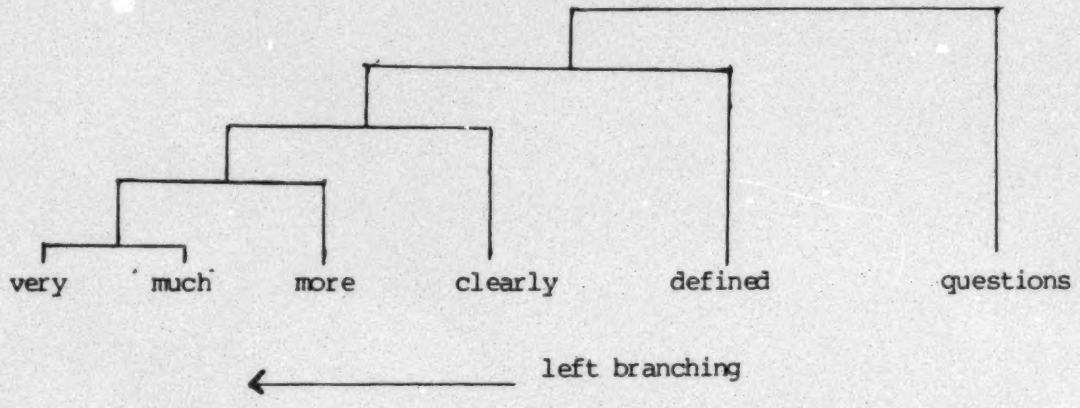
A further re-examination of the sentences in each composition revealed additional reasons for the lack of significant correlations between Williams' metric and the analytical ratings. These L2 compositions exhibited similarities to certain malfunctions seen in the compositions of basic writing (BW) students. There were instances in which the L2 writers attempted a complicated syntax for which they lacked the requisite competence (cf Bartholomae 1980); there were still other cases in which the L2 writers were not sure of what the written code allows for (cf Shaughnessy 1977:44). And furthermore, there

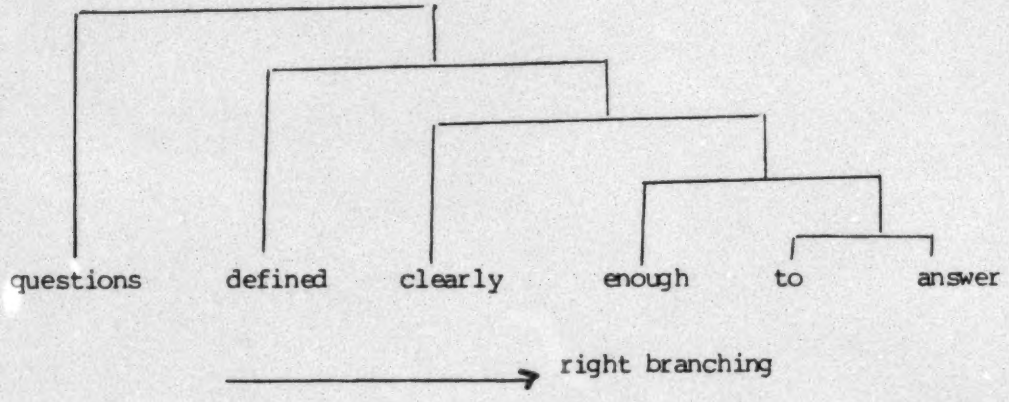
were breakdowns in some sentences in the L2 compositions which were similar to errors made by BW writers and described by Shaughnessy (1977: 47, 49, 52 respectively) as follows: "An erroneous form sets up different syntactic constraints from the ones the writer is obeying;" "patterns that erroneously combine features from several patterns, creating a kind of syntactic dissonance;" consolidation errors caused by "the writer, in his effort (or hurry) to compress, leav(ing) out a vital part of his statement." All the previously mentioned factors would account for the lack of significant correlations.

In sum, there is evidence that Williams' metric -- based on the assumption that the clearest style is that for which the grammatical structures redundantly support the perceived semantic structure -- has concurrent validity for L2 compositions. Though none of the rank order correlations reached significance, there were moderate correlations with analytical ratings of fluency, which assessed readability, and grammar which assessed cohesion, sentence structures, and sentence boundaries. The failure of the correlations to reach the criterion level of significance is attributed to writing malfunctions similar to those identified by Shaughnessy in BW.

FIGURE 1

AN EXAMPLE OF YNGVE'S DEPTH HYPOTHESIS
AS CITED BY WILLIAMS





(Williams 1979:599)

FIGURE 2

UNDERLYING REPRESENTATION

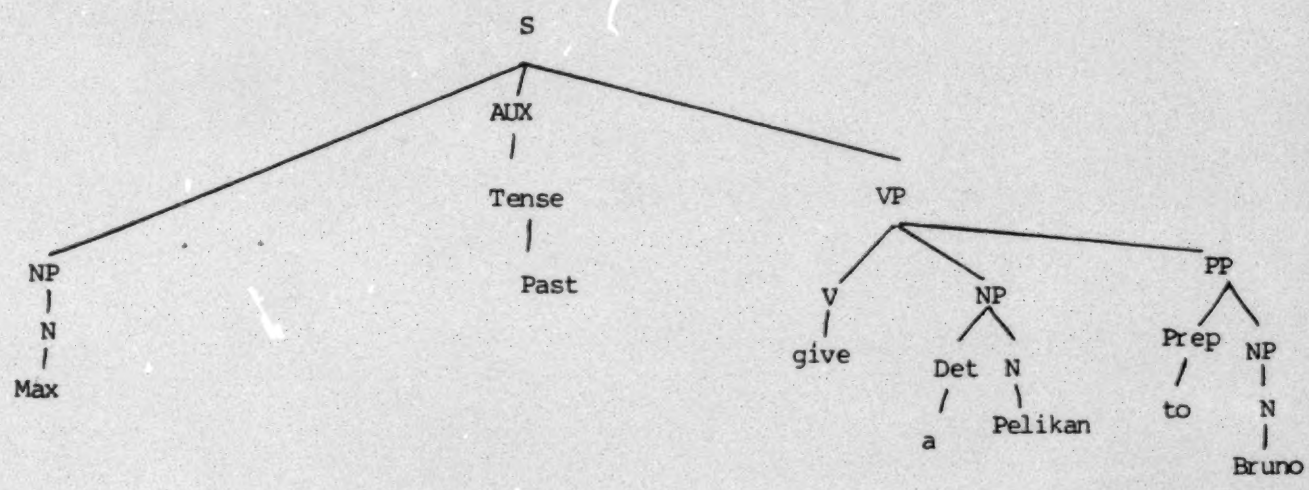


FIGURE 3
UNDERLYING REPRESENTATIONS

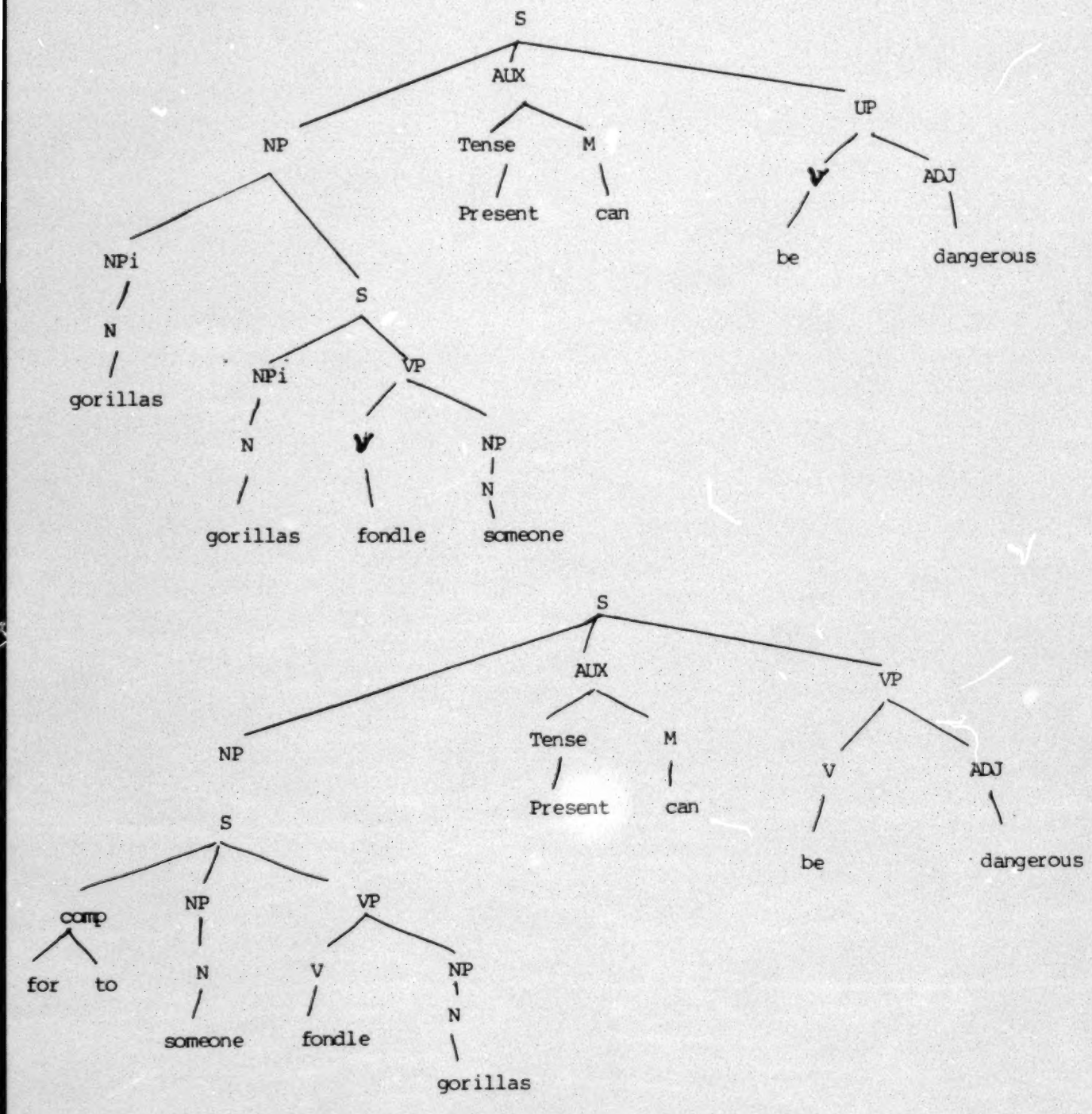


TABLE 1

WILLIAMS' COMPLEXITY METRIC AND ANALYTICAL RATINGS
SPEARMAN RHO'S

Fluency	.6750
Relevance	.5500
Grammar	.5417
Mechanics	.4667
Vocabulary	.2042

REFERENCES

Bartholomae, David (1980): "The study of error," College Composition and Cummunication XXXI, 253-269.

Bormuth, John R (1968): "New measures of grammatical complexity," in The Psycholinguistic Nature of the Reading Process, ed. by Kenneth S. Goodman, 239-253. Detroit: Wayne State University Press.

Clark, Herbert H (1965): "Some structural properties of simple active and passive sentences," Journal of Verbal Learning and Verbal Behavior 4, 365-370.

Fillmore, Charles (1968): "The case for case," in Universals in Linguistic Theory, ed. by Emmon Bach and Robert Harms. New York: Holt, Rinehart and Winston.

Flesch, Rudolf (1948): "A new readability yardstick," Journal of Applied Psychology 32, 221-233.

Flesch, Rudolf (1962a): The Art of Plain Talk. New York: Harper and Row.

Flesch, Rudolf (1962b): The Art of Readable Writing. New York: Harper and Row.

Fodor, J.A., T.G. Bever and M.F. Garrett (1974): The Psychology of Language. New York: McGraw-Hill.

Gibson, Walker (1966): Tough, Sweet and Stuffy. Bloomington, Indiana: Indiana University Press.

Hirsch, E.D., Jr (1977): The Philosophy of Composition. Chicago: The University of Chicago Press.

Hunt, Kellogg (1965): Grammatical Structures Written at Three Grade Levels, Research Report no. 3. Urbana, Illinois: National Council of Teachers of English.

Hunt, Kellogg (1970): "Syntactic maturity in school children and adults," Monographs of the Society for Reserach in Child Development 53, 1

Irmscher, William F (1979): Teaching Expository Writing. New York: Holt, Rinehart and Winston.

Johnson, M.G. (1967): "Syntactic position and rated meaning," Journal of Verbal Learning and Verbal Behavior 6, 240-246.

Mehler, J. (1963): "Some effects of grammatical transformations on the recall of English sentences," Journal of Verbal Learning and Verbal Behavior 2, 346-351.

Schlesinger, I.M. (1977): Production and Comprehension of Utterances. Hillsdale, New Jersey: Lawrence Erlbaum.

Shaughnessy, Mina P (1977): Errors and Expectations. A Guide for the Teacher of Basic Writing. New York: Oxford University Press.

Williams, Joseph M (1979): "Defining complexity," College English 40, 595-609.

Winterowd, W. Ross (1975): The Contemporary Writer. New York: Harcourt Brace Jovanovich, Inc.

Yngvé, Victor H (1960): "A model and hypothesis for language structure," Proceedings of the American Philosophical Society 104, 444-466.